Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Linear Algebra and its Applications 401 (2005) 159–172
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Inequalities for quantum relative entropy

N. Bebiano [a,*], R. Lemos [b,*], J. da Providência [c]

[a] *Department of Mathematics, University of Coimbra, P 3001-454 Coimbra, Portugal*
[b] *Department of Mathematics, University of Aveiro, P 3810-193 Aveiro, Portugal*
[c] *Department of Physics, University of Coimbra, P 3004-516 Coimbra, Portugal*

Received 27 October 2003; accepted 8 March 2004
Available online 25 June 2004
Submitted by R. Merris

Dedicated to G. de Oliveira

Abstract

Some logarithmic trace inequalities involving the notions of relative entropy are reobtained from a log-majorization result. The thermodynamic inequality is generalized and a chain of equivalent statements involving this inequality and the Peierls–Bogoliubov inequality is obtained.

Keywords: Relative entropy; Logarithmic trace inequalities; Peierls–Bogoliubov inequality; Thermodynamic inequality

1. Entropy

The concept of entropy was introduced in thermodynamics by Clausius in 1865, and some of the main steps towards the consolidation of the concept were taken by Boltzmann and Gibbs. Many generalizations and reformulations of this notion have been proposed, with motivations and applications in different subjects, such as statistical mechanics, information theory, dynamical systems and ergodic theory, biology, economics, human and social sciences.

In quantum mechanics, pure states of physical systems are described by vectors in a Hilbert space, while mixed states are described by positive semi-definite

* Corresponding authors.
E-mail addresses: bebiano@ci.uc.pt (N. Bebiano), rute@mat.ua.pt (R. Lemos), providencia@teor.fis.uc.pt (J. da Providência).

0024-3795/$ - see front matter
doi:10.1016/j.laa.2004.03.023

matrices with trace one. Such matrices are called *density matrices*. The eigenvalues of a density matrix are the *probabilities* that the system under consideration is in the pure states described by the associated eigenvectors.

In classical commutative systems, physical states can be identified with diagonal density matrices, which are associated with *probability vectors* $p = (p_1, \ldots, p_n)$, that is, $p_i \geqslant 0$, $i = 1, \ldots, n$, and $\sum_{i=1}^n p_i = 1$. In this context, the *Shannon entropy* is defined by

$$S(p) = -\sum_{i=1}^{n} p_i \log p_i$$

(convention: $x \log x = 0$ if $x = 0$). Claude Shannon, in his pioneer work in information theory [17], considered the entropy from an axiomatic point of view, regarding it as the measure of efficiency of a communication system. That terminology was suggested by von Neumann, due to the analogy between the concepts in physics and in information theory.

In quantum systems, the *entropy* of a mixed state described by the density matrix A is defined by

$$S(A) = -\mathrm{Tr}(A \log A).$$

This notion was introduced by von Neumann as the degree of disorder of the system. The quantum entropy is clearly invariant under unitary similarity transformations, that is, $S(U^*AU) = S(A)$, for all unitary matrices U. Assuming that λ_i, $i = 1, \ldots, n$, are the eigenvalues of the density matrix A, the entropy of A is given by

$$S(A) = -\sum_{i=1}^{n} \lambda_i \log \lambda_i.$$

The minimum entropy $S(A) = 0$ occurs if and only if one of the eigenvalues of the density matrix A is one and all the others are zero. The maximum entropy $S(A) = \log n$ occurs if and only if all the eigenvalues of A are equal, that is, $A = I_n/n$.

The *relative entropy* of two states of a commutative system, described by probability vectors p and q, is called the *Kullback–Leibler relative entropy* or *information divergence* [13], and is defined by

$$S(p, q) = \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i}$$

(convention: $x \log x = 0$ if $x = 0$, and $x \log y = +\infty$ if $y = 0$ and $x \neq 0$).

Its non-commutative or quantum analogue, for density matrices A and B, is the *Umegaki relative entropy* [19], defined by

$$S(A, B) = \mathrm{Tr}(A(\log A - \log B)).$$

Since the condition of unital trace is not essential, sometimes we relax it. Clearly, $S(A) = -S(A, I_n)$.

In the sequel, the notions of entropy will be considered in the general setup of $n \times n$ positive semi-definite matrices. We write $A \geqslant 0$ if A is a positive semi-definite matrix. If $A \geqslant 0$ is invertible, we write $A > 0$.

The *relative operator entropy* of $A, B > 0$ was introduced in noncomutative information theory by Fujii and Kamei [7], and is defined by

$$\hat{S}(A|B) = A^{1/2} \log(A^{-1/2} B A^{-1/2}) A^{1/2}.$$

Previously, Nakamura and Umegaki [16] had introduced the operator entropy $-A \log A$ and the quantity $-\mathrm{Tr}\, \hat{S}(A|B)$ had already been discussed by Belavkin and Staszewski [6], in the general setup of physical systems described by C^*-algebras.

The relative entropy satisfies $S(U^*AU, U^*BU) = S(A, B)$, for all unitary matrices U. If A and B commute, they are simultaneously unitarily diagonalizable and so

$$-\mathrm{Tr}\, \hat{S}(A|B) = S(A, B) = \sum_{i=1}^{n} \lambda_i \log \frac{\lambda_i}{\gamma_i},$$

where λ_i and γ_i, $i = 1, \ldots, n$, are the eigenvalues (with simultaneous eigenvectors) of A and B, respectively.

This note is organized as follows. In Section 2, a log-majorization result is presented and some logarithmic trace inequalities involving the notions of relative entropy are reobtained. In Section 3, the Peierls–Bogoliubov inequality is generalized. In Section 4, the thermodynamic inequality is generalized and a chain of equivalent statements involving this inequality and the Peierls–Bogoliubov inequality is obtained.

2. Logarithmic trace inequalities

For a Hermitian matrix H in M_n, the algebra of $n \times n$ complex matrices, we assume that the eigenvalues $\lambda_i(H)$, $i = 1, \ldots, n$, are arranged in non-increasing order $\lambda_1(H) \geqslant \cdots \geqslant \lambda_n(H)$. For Hermitian matrices $A, B \geqslant 0$, the *log-majorization* of A by B, denoted by $A \prec_{(\log)} B$, is defined as

$$\prod_{i=1}^{k} \lambda_i(A) \leqslant \prod_{i=1}^{k} \lambda_i(B), \quad k = 1, \ldots, n, \tag{1}$$

and $\det(A) = \det(B)$. If $A \geqslant 0$, then $\lambda_1(A^{(k)}) = \prod_{i=1}^{k} \lambda_i(A)$, where $A^{(k)}$ denotes the kth compound of the matrix A [15]. Thus, (1) can be written as $\lambda_1(A^{(k)}) \leqslant \lambda_1(B^{(k)})$, $k = 1, \ldots, n$. It is well-known that the log-majorization $A \prec_{(\log)} B$ implies the *weak majorization* $A \prec_w B$, that is,

$$\sum_{i=1}^{k} \lambda_i(A) \leqslant \sum_{i=1}^{k} \lambda_i(B), \quad k = 1, \ldots, n.$$

This concept is the basis of a powerful technique for deriving matrix norm inequalities $\|A\| \leqslant \|B\|$, for any unitarily invariant norm $\|\cdot\|$ and, in particular, trace inequalities.

For example, Araki [3] obtained, as an extension of the Lieb–Thirring trace inequality, the following log-majorization

$$(A^{1/2}BA^{1/2})^s \prec_{(\log)} A^{s/2}B^sA^{s/2}, \quad s \geqslant 1.$$

This is equivalent to

$$(A^{q/2}B^qA^{q/2})^{1/q} \prec_{(\log)} (A^{p/2}B^pA^{p/2})^{1/p}, \quad 0 < q \leqslant p.$$

Using techniques of Ando and Hiai [1], a log-majorization of this type is obtained in Theorem 2.1. A fundamental tool in the proof of Theorem 2.1 is the *Furuta inequality* [8], which asserts that if $X \geqslant Y \geqslant 0$, then

$$(X^{s/2}X^rX^{s/2})^\alpha \geqslant (X^{s/2}Y^rX^{s/2})^\alpha,$$

for $r, s \geqslant 0$ and $0 \leqslant \alpha \leqslant 1$ such that $(r+s)\,\alpha \leqslant 1+s$.

Theorem 2.1. *If* $A, B \geqslant 0$, *then*

$$A^{(1+q)/2}B^qA^{(1+q)/2} \prec_{(\log)} A^{1/2}(A^{p/2}B^pA^{p/2})^{q/p}A^{1/2}, \quad 0 < q \leqslant p. \tag{2}$$

Proof. We show that

$$\lambda_1(A^{(1+q)/2}B^qA^{(1+q)/2}) \leqslant \lambda_1(A^{1/2}(A^{p/2}B^pA^{p/2})^{q/p}A^{1/2}), \quad 0 < q \leqslant p. \tag{3}$$

Suppose that there exists $\gamma > 0$ such that

$$\lambda_1(A^{(1+q)/2}B^qA^{(1+q)/2}) > \gamma \geqslant \lambda_1(A^{1/2}(A^{p/2}B^pA^{p/2})^{q/p}A^{1/2}). \tag{4}$$

Taking $C = \gamma^{1/(1+q)}A$, $D = \gamma^{-2/q}B$, dividing (4) by γ, and bearing in mind that

$$\lambda_1(C^{(1+q)/2}D^qC^{(1+q)/2}) = \gamma^{-1}\,\lambda_1(A^{(1+q)/2}B^qA^{(1+q)/2}),$$
$$\lambda_1(C^{1/2}(C^{p/2}D^pC^{p/2})^{q/p}C^{1/2}) = \gamma^{-1}\,\lambda_1(A^{1/2}(A^{p/2}B^pA^{p/2})^{q/p}A^{1/2}),$$

we have

$$\lambda_1(C^{(1+q)/2}D^qC^{(1+q)/2}) > 1 \geqslant \lambda_1(C^{1/2}(C^{p/2}D^pC^{p/2})^{q/p}C^{1/2}). \tag{5}$$

From (5), we get $I_n \geqslant C^{1/2}(C^{p/2}D^pC^{p/2})^{q/p}C^{1/2}$, which implies

$$C^{-1} \geqslant (C^{p/2}D^pC^{p/2})^{q/p} \geqslant 0.$$

By Furuta's inequality, with $\alpha = q/p$, $r = p/q$ and $s = p$, we have $C^{-1-q} \geqslant D^q$. Hence, $I_n \geqslant C^{(1+q)/2}D^qC^{(1+q)/2}$ and so $1 \geqslant \lambda_1(C^{(1+q)/2}D^qC^{(1+q)/2})$, contradicting (5). Thus, (3) is proved. Now, replacing in (3) A and B by $A^{(k)}$ and $B^{(k)}$, $k = 1, \ldots, n$, respectively, using the Binet–Cauchy identity $A^{(k)}B^{(k)} = (AB)^{(k)}$, and observing that the determinants of both sides of (2) are equal, the result follows. □

Using Theorem 2.1, we give a new proof to the following result due to Hiai and Petz [10], and lately strengthened by Ando and Hiai [1].

Theorem 2.2. *If* $A, B \geqslant 0$, *then*

$$\mathrm{Tr}(A(\log A + \log B)) \leqslant \frac{1}{p}\mathrm{Tr}(A \log(A^{p/2} B^p A^{p/2})), \quad p > 0, \tag{6}$$

and the right hand side of (6) *converges decreasingly to the left hand side as* $p \downarrow 0$.

Proof. We may assume $B > 0$ (cf. [10], Theorem 3.5). We consider $A > 0$ due to the continuity of $\mathrm{Tr}(A \log(A^{p/2} B^p A^{p/2}))$ in $A \geqslant 0$. Since log-majorization implies weak majorization, from Theorem 2.1, we have

$$\mathrm{Tr}(A^{1+q} B^q) \leqslant \mathrm{Tr}(A(A^{p/2} B^p A^{p/2})^{q/p}), \quad 0 < q \leqslant p,$$

with both sides equal to $\mathrm{Tr}(A)$ when $q = 0$. So

$$\frac{\mathrm{d}}{\mathrm{d}q}\mathrm{Tr}(A^{1+q} B^q)\bigg|_{q=0} \leqslant \frac{\mathrm{d}}{\mathrm{d}q}\mathrm{Tr}\,(A(A^{p/2} B^p A^{p/2})^{q/p})\bigg|_{q=0}, \quad p > 0,$$

and (6) follows from the computation of these derivatives.

By standard arguments, $\log(A^{p/2} B^p A^{p/2})$ extends to an analytic function in some neighborhood of the origin. Straightforward computations lead to the following power series expansion (cf. [11], Theorem 4.1)

$$\begin{aligned}\log(A^{p/2} B^p A^{p/2}) = {} & p(\log A + \log B) \\ & + \frac{p^3}{24}\big[[\log A, \log B], (2 \log B + \log A)\big] + \cdots,\end{aligned}$$

which holds for any $p \in \mathbb{R}$ in a certain neighborhood of 0. (Here, the usual notation for the commutator of matrices $[X, Y] = XY - YX$ is used.) Since $[A, \log A] = 0$, we may easily conclude that

$$\mathrm{Tr}\left(A\big[[\log A, \log B], (2 \log B + \log A)\big]\right) = 4\mathrm{Tr}\,(A[\log A, \log B] \log B) \geqslant 0.$$

Thus, the right hand side of (6) decreases to the left hand side as $p \downarrow 0$. □

Hiai [11] proved that the equality occurs in the logarithmic trace inequality (6) if and only if $AB = BA$.

For $A, B > 0$, (6) can be rewritten in entropy terminology as

$$S(A, B) \leqslant -\frac{1}{p}\,\mathrm{Tr}(\hat{S}(A^p | B^p) A^{1-p}), \quad p > 0. \tag{7}$$

For $p = 1$, (7) establishes a relation between the Umegaki relative entropy and the version due to Belavkin and Staszewski: $S(A, B) \leqslant -\mathrm{Tr}\,\hat{S}(A|B)$.

For $0 \leqslant \alpha \leqslant 1$, the *α-power mean* of matrices $A > 0$, $B \geqslant 0$, is defined by

$$A \#_\alpha B = A^{1/2}(A^{-1/2} B A^{-1/2})^\alpha A^{1/2}.$$

If A is not strictly positive, then the α-*power mean* of $A, B \geqslant 0$ is

$$A\#_\alpha B = \lim_{\epsilon\downarrow 0}(A + \epsilon I)\#_\alpha B.$$

In the Kubo–Ando theory of operator means [12], the α-power mean is the one corresponding to the operator monotone function x^α. For $0 \leqslant \alpha \leqslant 1$, we have $A\#_\alpha B \geqslant 0$. In particular, $A\#_0 B = A$, $A\#_1 B = B$ and $A\#B = A\#_{1/2}B$ is the *geometric mean* of A, B (this terminology is due to the fact that if A and B commute, then $A\#B = (AB)^{1/2}$). The following equality holds

$$\frac{\mathrm{d}}{\mathrm{d}\alpha} A\#_\alpha B\bigg|_{\alpha=0} = \hat{S}(A|B). \tag{8}$$

Corollary 2.1. *If* $A, B > 0$ *and* $0 \leqslant \alpha \leqslant 1$, *then*

$$S(A, (A^r\#_\alpha B^s)^{t/p}) \leqslant -\frac{1}{p}\,\mathrm{Tr}(\hat{S}(A^p|\,(A^r\#_\alpha B^s)^t)A^{1-p}),$$

for $r, s \geqslant 0$, $t \geqslant 1$ *and* $p > 0$.

Proof. Replacing B by $(A^r\#_\alpha B^s)^{t/p}$ in (7), the result follows. $\square$

Note that (7) is recovered by Corollary 2.1 in the particular case $\alpha = 1$, $p = st$ and $s \neq 0$.

Using Corollary 2.1, the continuous parameter version of *Lie–Trotter's formula*

$$\mathrm{e}^{X+Y} = \lim_{p\to 0}(\mathrm{e}^{pX/2}\mathrm{e}^{pY}\mathrm{e}^{pX/2})^{1/p}, \quad X, Y \in M_n,$$

and its α-power mean variant [10]:

$$\mathrm{e}^{(1-\alpha)X+\alpha Y} = \lim_{p\to 0}(\mathrm{e}^{pX}\#_\alpha \mathrm{e}^{pY})^{1/p}, \quad X, Y \in M_n, \tag{9}$$

we give in Corollary 2.2 an alternative proof for the following logarithmic trace inequality of Ando and Hiai [1].

Corollary 2.2. *If* $A \geqslant 0$, $B > 0$, $0 \leqslant \alpha \leqslant 1$ *and* $p > 0$, *then*

$$\frac{1}{p}\,\mathrm{Tr}(A\log(A^p\#_\alpha B^p)) + \frac{\alpha}{p}\mathrm{Tr}(A\log(A^{p/2}B^{-p}A^{p/2})) \geqslant \mathrm{Tr}(A\log A),$$

and the left hand side converges to $\mathrm{Tr}(A\log A)$ *as* $p\downarrow 0$.

Proof. Firstly, let $A > 0$, and consider Corollary 2.1 in the particular case $t = 1$ and $p = r = s$. Since

$$S(A, (A^p\#_\alpha B^p)^{1/p}) = \mathrm{Tr}(A\log A) - \frac{1}{p}\,\mathrm{Tr}(A\log(A^p\#_\alpha B^p)),$$

$$\mathrm{Tr}(\hat{S}(A^p|A^p\#_\alpha B^p)A^{1-p}) = -\alpha\,\mathrm{Tr}(A\log(A^{p/2}B^{-p}A^{p/2})),$$

the asserted inequality follows. Now, we study the convergence as $p \downarrow 0$. Considering (9) for the Hermitian matrices $X = \log A$ and $Y = \log B$, and having in mind that log is a continuous function, we get

$$\lim_{p\downarrow 0} \log(A^p \#_\alpha B^p)^{1/p} = (1-\alpha)\log A + \alpha \log B. \tag{10}$$

On the other hand, by the parameter version of Lie–Trotter's formula applied to the matrices $X = \log A$ and $Y = -\log B$, we obtain:

$$\lim_{p\downarrow 0} \log(A^{p/2} B^{-p} A^{p/2})^{1/p} = \log A - \log B. \tag{11}$$

The corollary easily follows from (10) and (11). If $A \geqslant 0$, by a perturbation, we take $A + \epsilon I_n$, $\epsilon > 0$, and the result follows by a continuity argument. $\square$

In relative entropy terminology, Corollary 2.2 establishes that

$$S(A, (A^p \#_\alpha B^p)^{1/p}) \leqslant -\frac{\alpha}{p} \operatorname{Tr}(\hat{S}(A^p|B^p)A^{1-p}). \tag{12}$$

The proof of Corollary 2.2 shows that the right and the left hand side of (12) converge to $\alpha S(A, B)$ as $p \downarrow 0$.

3. On the Peierls–Bogoliubov inequality

It is an important issue in statistical mechanics to calculate the value of the so-called *partition function* $\operatorname{Tr}(e^{\hat{H}})$, where the Hermitian matrix $\hat{H}$ is the Hamiltonian of a physical system. Since that computation is often difficult, it is simpler to compute the related quantity $\operatorname{Tr}(e^H)$, where H is a convenient approximation of the Hamiltonian $\hat{H}$. Indeed, let $\hat{H} = H + K$. The *Peierls–Bogoliubov inequality* provides useful information on $\operatorname{Tr}(e^{H+K})$ from $\operatorname{Tr}(e^H)$. This inequality states that, for two Hermitian operators H and K,

$$\operatorname{Tr}(e^H) \exp \frac{\operatorname{Tr}(e^H K)}{\operatorname{Tr}(e^H)} \leqslant \operatorname{Tr}(e^{H+K}).$$

The equality occurs in the Peierls–Bogoliubov inequality if and only if K is a scalar matrix. This well-known inequality will be extended in Corollary 3.3 of the following theorem of Ando and Hiai [2].

Theorem 3.1. *If $A_1, B_1 \geqslant 0$ and $A_2, B_2 > 0$, then*

$$\begin{aligned}\operatorname{Tr}(A_1B_1)\, \log \frac{\operatorname{Tr}(A_1B_1)}{\operatorname{Tr}(A_2B_2)} &\leqslant \operatorname{Tr}\left(A_1^{1/2}B_1A_1^{1/2} \log\left(A_1^{1/2}B_2^{-1}A_1^{1/2}\right)\right)\\ &\quad + \operatorname{Tr}\left(B_1^{1/2}A_1B_1^{1/2} \log\left(B_1^{1/2}A_2^{-1}B_1^{1/2}\right)\right).\end{aligned} \tag{13}$$

Relaxing the condition of unital trace in the definition of relative entropy, (13) can be written in the condensed form:

$$\frac{1}{n}S(\mathrm{Tr}(A_1B_1)I_n, \mathrm{Tr}(A_2B_2)I_n) \leqslant -\mathrm{Tr}(\hat{S}(A_1|B_2)B_1) - \mathrm{Tr}(\hat{S}(B_1|A_2)A_1). \tag{14}$$

Corollary 3.1. *If* $A_1, B_1 \geqslant 0$ *and* $A_2, B_2 > 0$, *then*

$$\frac{1}{n}S(\mathrm{Tr}(A_1B_1)\,I_n, \mathrm{Tr}((A_1\#_\alpha A_2)\,(B_1\#_\beta B_2))\,I_n) \leqslant -\alpha\,\mathrm{Tr}(\hat{S}(A_1|A_2)B_1) - \beta\,\mathrm{Tr}(\hat{S}(B_1|B_2)A_1),$$

for every $0 \leqslant \alpha \leqslant 1$ *and* $0 \leqslant \beta \leqslant 1$.

Proof. Replacing B_2, A_2 in (14) by $A_1\#_\alpha A_2$, $B_1\#_\beta B_2$, respectively, and observing that $\mathrm{Tr}(\hat{S}(A_1|A_1\#_\alpha A_2)B_1) = \alpha\,\mathrm{Tr}(\hat{S}(A_1|A_2)B_1)$, $\mathrm{Tr}(\hat{S}(B_1|B_1\#_\beta B_2)A_1) = \beta\,\mathrm{Tr}(\hat{S}(B_1|B_2)A_1)$, the result follows. □

Corollary 3.2. *If* $A_1, B_1 \geqslant 0$ *and* $A_2, B_2 > 0$, *then the function*

$$f(\alpha) = \log \mathrm{Tr}((A_1\#_\alpha A_2)(B_1\#_\alpha B_2)), \quad 0 \leqslant \alpha \leqslant 1,$$

is such that $f'(0) \leqslant f(1) - f(0)$.

Proof. Let $A_1, B_1 \geqslant 0$ and $A_2, B_2 > 0$, and interchange in (14) the roles of the matrices A_2 and B_2. Dividing the left and the right hand side of the resulting inequality by $-\mathrm{Tr}(A_1B_1)$, taking the exponential, and multiplying both sides of the so-obtained inequality by $\mathrm{Tr}(A_1B_1)$, we write (14) in the form

$$\mathrm{Tr}(A_1B_1)\exp\frac{\mathrm{Tr}(\hat{S}(A_1|A_2)B_1) + \mathrm{Tr}(\hat{S}(B_1|B_2)A_1)}{\mathrm{Tr}(A_1B_1)} \leqslant \mathrm{Tr}(A_2B_2). \tag{15}$$

From the definition of f, it is clear that $\mathrm{e}^{f(0)} = \mathrm{Tr}(A_1B_1)$ and $\mathrm{e}^{f(1)} = \mathrm{Tr}(A_2B_2)$.

On the other hand, taking the derivative of f with respect to α at the origin and recalling (8), we obtain

$$f'(0) = \frac{\mathrm{Tr}(\hat{S}(A_1|A_2)B_1 + A_1\hat{S}(B_1|\,B_2))}{\mathrm{Tr}(A_1B_1)}.$$

Thus, we can rewrite (15) in the form $\mathrm{e}^{f(0)}\mathrm{e}^{f'(0)} \leqslant \mathrm{e}^{f(1)}$ and the result easily follows. □

Corollary 3.3. *For Hermitian matrices* G, H, K *and* L, *we have*

$$\mathrm{Tr}(\mathrm{e}^H\mathrm{e}^G)\exp\frac{\mathrm{Tr}(\hat{S}(\mathrm{e}^H|\,\mathrm{e}^{G+L})\,\mathrm{e}^G) + \mathrm{Tr}(\hat{S}(\mathrm{e}^G|\,\mathrm{e}^{H+K})\,\mathrm{e}^H)}{\mathrm{Tr}(\mathrm{e}^H\mathrm{e}^G)} \leqslant \mathrm{Tr}(\mathrm{e}^{H+K}\mathrm{e}^{G+L}).$$

Proof. Considering $A_1 = \mathrm{e}^H$, $A_2 = \mathrm{e}^{G+L}$, $B_1 = \mathrm{e}^G$ and $B_2 = \mathrm{e}^{H+K}$ in (15) the result follows. □

Remarks. Theorem 3.1 is recovered by Corollary 3.1 for $\alpha = \beta = 1$. Corollary 3.3 is an extension of the Peierls–Bogoliubov inequality, which is recovered in the particular case $G = L = 0$. This can be easily seen observing that

$$\operatorname{Tr} \hat{S}(\mathrm{e}^H \,|\, I_n) = -\operatorname{Tr}(\mathrm{e}^H H) \quad \text{and} \quad \operatorname{Tr}(\hat{S}(I_n \,|\, \mathrm{e}^{H+K})\, \mathrm{e}^H) = \operatorname{Tr}(\mathrm{e}^H (H + K)).$$

4. The thermodynamic inequality and equivalent statements

The quantum observables are modeled by Hermitian matrices. For example, the energy operator H is a Hermitian operator. For H the energy operator, the *statistical energy mean value* of the state described by the density matrix A is

$$E = \operatorname{Tr}(AH)$$

and the *free energy* of that state is

$$\psi(A) = \operatorname{Tr}(AH) - \theta\, S(A),$$

where $\theta = kT$, k is the Boltzmann constant and T is the absolute temperature. The evaluation of the extremum of $\psi(A)$ is an important problem in physics. For convenience, we consider $\theta = -1$.

Theorem 4.1 [5]. *If H is a Hermitian matrix, then*

$$\log \operatorname{Tr}(\mathrm{e}^H) = \max\{\operatorname{Tr}(AH) + S(A) : A \geqslant 0,\ \operatorname{Tr}(A) = 1\}.$$

Theorem 4.1 implies the important *thermodynamic inequality* [4,14]:

$$\log \operatorname{Tr}(\mathrm{e}^H) \geqslant \operatorname{Tr}(AH) + S(A), \tag{16}$$

where the equality occurs if and only if $A = \mathrm{e}^H / \operatorname{Tr}(\mathrm{e}^H)$. As observed by the Referee, the inequality (16) and its equality case are immediate consequences of the strict positivity of the relative entropy: $S(A, \mathrm{e}^H/\operatorname{Tr}(\mathrm{e}^H)) \geqslant 0$.

Replacing H by $-H/\theta$ in (16) and multiplying both members by $-\theta$, we conclude that $\psi(A)$ is an approximation (an upper bound) to the *Helmholtz free energy function* $F = -\theta \log \operatorname{Tr}(\mathrm{e}^{-H/\theta})$.

On the other hand, replacing H in Theorem 4.1 by $H + \log B$, $B > 0$, we obtain the following result of Hiai and Petz. Clearly, Theorem 4.1 is a particular case of Corollary 4.1 when $B = I_n$.

Corollary 4.1 [10]. *If $B > 0$ and H is a Hermitian matrix, then*

$$\log \operatorname{Tr}(\mathrm{e}^{H+\log B}) = \max\{\operatorname{Tr}(AH) - S(A, B) : A \geqslant 0,\ \operatorname{Tr}(A) = 1\}.$$

For Hermitian matrices H and K, the famous *Golden–Thompson inequality* holds: $\operatorname{Tr}(\mathrm{e}^H \mathrm{e}^K) \geqslant \operatorname{Tr}(\mathrm{e}^{H+K})$. This is one of the earlier trace inequalities [9,18] in statistical mechanics. Using this inequality, a generalization of the thermodynamic inequality is obtained in Theorem 4.2.

Theorem 4.2. *Let* $A, B > 0$, *with* $\mathrm{Tr}(A) = 1$ *and let* H *be Hermitian. Then*

$$\log \mathrm{Tr}(\mathrm{e}^H (A^p \#_\alpha B^p)^{1/p}) \geqslant \mathrm{Tr}(AH) + \frac{\alpha}{p} \mathrm{Tr}(\hat{S}(A^p|B^p)A^{1-p}), \tag{17}$$

for $0 \leqslant \alpha \leqslant 1$ *and* $p > 0$. *If* $\alpha = 1$ *and* $B = I_n$, (17) *reduces to the thermodynamic inequality.*

Proof. If A, $B > 0$, then $(A^p \#_\alpha B^p)^{1/p} > 0$, for $0 \leqslant \alpha \leqslant 1$ and $p > 0$. By the Golden–Thompson inequality applied to the Hermitian matrices H and $\log(A^p \#_\alpha B^p)^{1/p}$, and by the monotonicity of the logarithm, we have

$$\log \mathrm{Tr}(\mathrm{e}^H (A^p \#_\alpha B^p)^{1/p}) \geqslant \log \mathrm{Tr}(\mathrm{e}^{H+\log(A^p \#_\alpha B^p)^{1/p}}), \quad 0 \leqslant \alpha \leqslant 1, \; p > 0.$$

Replacing B by $(A^p \#_\alpha B^p)^{1/p}$ in Corollary 4.1, we get

$$\log \mathrm{Tr}(\mathrm{e}^{H+\log(A^p \#_\alpha B^p)^{1/p}}) \geqslant \mathrm{Tr}(AH) - S(A, (A^p \#_\alpha B^p)^{1/p}), \quad 0 \leqslant \alpha \leqslant 1, \; p > 0.$$

These inequalities in conjunction with the logarithmic trace inequality (12) yield (17). To finish the proof, let $\alpha = 1$. We obtain

$$\log \mathrm{Tr}(\mathrm{e}^H B) \geqslant \mathrm{Tr}(AH) + \frac{1}{p} \mathrm{Tr}(\hat{S}(A^p|B^p)A^{1-p}), \quad p > 0.$$

Moreover, if $B = I_n$ then $\mathrm{Tr}(\hat{S}(A^p|B^p)A^{1-p}) = p\, S(A)$, and the thermodynamic inequality is obtained. □

We present now two equivalent statements to Theorem 4.2.

Theorem 4.3. *For* $0 \leqslant \alpha \leqslant 1$ *and* $p > 0$, *the following conditions are equivalent*:

(i) *Let* $A, B > 0$, *with* $\mathrm{Tr}(A) = 1$ *and* H *be Hermitian. Then*

$$\log \mathrm{Tr}(\mathrm{e}^H (A^p \#_\alpha B^p)^{1/p}) \geqslant \mathrm{Tr}(AH) + \frac{\alpha}{p} \mathrm{Tr}(\hat{S}(A^p|B^p)A^{1-p}).$$

(ii) *If* $A_1, A_2, B_2 > 0$, *then*

$$\frac{1}{n} S(\mathrm{Tr}(A_1)\, I_n, \mathrm{Tr}(B_2(A_1^p \#_\alpha A_2^p)^{1/p})\, I_n) \leqslant -\frac{\alpha}{p} \mathrm{Tr}\left(\hat{S}\left(A_1^p|A_2^p\right) A_1^{1-p}\right) - \mathrm{Tr}(\hat{S}(I_n|B_2)A_1).$$

(iii) *For Hermitian matrices* H, K, L, *we have*

$$\mathrm{Tr}(\mathrm{e}^H) \exp \frac{p\, \mathrm{Tr}(\mathrm{e}^H (H+K)) + \alpha\, \mathrm{Tr}(\hat{S}(\mathrm{e}^{pH}|\mathrm{e}^{pL})\, \mathrm{e}^{(1-p)H})}{p\, \mathrm{Tr}(\mathrm{e}^H)} \leqslant \mathrm{Tr}(\mathrm{e}^{H+K} (\mathrm{e}^{pH} \#_\alpha \mathrm{e}^{pL})^{1/p}).$$

Proof. (i) $\Rightarrow$ (ii): Let $A_1,\ A_2, B_2 > 0$. Consider $A = A_1/\mathrm{Tr}(A_1)$, $B = A_2$ and $H = \log B_2$. These matrices satisfy the conditions in (i) and satisfy the following properties:

$$(A^p \#_\alpha B^p)^{1/p} = (\mathrm{Tr}\, A_1)^{\alpha-1}(A_1^p \#_\alpha A_2^p)^{1/p}, \tag{18}$$

$$\mathrm{Tr}(AH) = \frac{\mathrm{Tr}(\hat{S}(I_n|B_2)A_1)}{\mathrm{Tr}(A_1)}, \tag{19}$$

$$\mathrm{Tr}(\hat{S}(A^p|B^p)\, A^{1-p}) = \frac{\mathrm{Tr}(\hat{S}(A_1^p|A_2^p)A_1^{1-p})}{\mathrm{Tr}(A_1)} + p\ \log \mathrm{Tr}(A_1). \tag{20}$$

Replacing (18), (19) and (20) in the inequality in (i) and multiplying both members by $-\mathrm{Tr}(A_1)$, we obtain

$$\begin{aligned} &-\mathrm{Tr}(A_1)(\log(\mathrm{Tr}\, A_1)^{\alpha-1} + \log \mathrm{Tr}(B_2(A_1^p \#_\alpha A_2^p)^{1/p})) \\ &\quad \leqslant -\mathrm{Tr}(\hat{S}(I_n|B_2)A_1) - \frac{\alpha}{p}\,\mathrm{Tr}(\hat{S}(A_1^p|A_2^p)A_1^{1-p}) \\ &\qquad -\alpha\,\mathrm{Tr}(A_1)\log \mathrm{Tr}(A_1). \end{aligned} \tag{21}$$

Summing the left hand side of (21) with $\alpha\,\mathrm{Tr}(A_1)\log\mathrm{Tr}(A_1)$, we get

$$\frac{1}{n}S(\mathrm{Tr}(A_1)\, I_n, \mathrm{Tr}(B_2(A_1^p \#_\alpha A_2^p)^{1/p})\, I_n)$$

and the implication is proved.
(ii) $\Rightarrow$ (iii): Let H, K, L be Hermitian matrices, and consider $A_1 = \mathrm{e}^H$, $A_2 = \mathrm{e}^L$ and $B_2 = \mathrm{e}^{H+K}$. Then

$$\begin{aligned} &\frac{1}{n}S\left(\mathrm{Tr}(A_1)I_n, \mathrm{Tr}\left(B_2(A_1^p \#_\alpha A_2^p)^{\frac{1}{p}}\right) I_n\right) \\ &\quad = -\mathrm{Tr}(\mathrm{e}^H)\log \frac{\mathrm{Tr}\left(\mathrm{e}^{H+K}(\mathrm{e}^{pH} \#_\alpha \mathrm{e}^{pL})^{\frac{1}{p}}\right)}{\mathrm{Tr}(\mathrm{e}^H)}, \end{aligned} \tag{22}$$

$$\mathrm{Tr}(\hat{S}(I_n|B_2)A_1) = \mathrm{Tr}(\mathrm{e}^H(H+K)). \tag{23}$$

Replacing (22) and (23) in the inequality in (ii) and dividing both members of the so obtained inequality by $-\mathrm{Tr}(\mathrm{e}^H)$, we obtain

$$\begin{aligned} \log \frac{\mathrm{Tr}(\mathrm{e}^{H+K}(\mathrm{e}^{pH} \#_\alpha \mathrm{e}^{pL})^{1/p})}{\mathrm{Tr}(\mathrm{e}^H)} \geqslant \frac{\alpha\, \mathrm{Tr}(\hat{S}(\mathrm{e}^{pH}|\mathrm{e}^{pL})\, \mathrm{e}^{(1-p)H})}{p\,\mathrm{Tr}(\mathrm{e}^H)} \\ + \frac{\mathrm{Tr}(\mathrm{e}^H(H+K))}{\mathrm{Tr}(\mathrm{e}^H)}. \end{aligned}$$

Taking the exponential, we easily find

$$\frac{\operatorname{Tr}(e^{H+K}(e^{pH}\#_\alpha e^{pL})^{1/p})}{\operatorname{Tr}(e^H)}$$
$$\geqslant \exp\frac{p\operatorname{Tr}(e^H(H+K))+\alpha\operatorname{Tr}(\hat{S}(e^{pH}|e^{pL})\,e^{(1-p)H})}{p\operatorname{Tr}(e^H)}.$$

(iii) $\Rightarrow$ (i): If $A, B > 0$, then there exist Hermitian matrices H and L such that $A = e^H$ and $B = e^L$. Since H is Hermitian and $K = H - \log A$ is also Hermitian, (iii) implies that

$$\operatorname{Tr}(A)\exp\frac{p\operatorname{Tr}(AH)+\alpha\operatorname{Tr}(\hat{S}(A^p|B^p)\,A^{(1-p)})}{p\operatorname{Tr}(A)} \leqslant \operatorname{Tr}(e^H(A^p\#_\alpha B^p)^{1/p}).$$

Since $\operatorname{Tr}(A) = 1$, the monotonicity of the logarithmic function yields

$$\operatorname{Tr}(AH)+\frac{\alpha}{p}\operatorname{Tr}(\hat{S}(A^p|B^p)\,A^{1-p}) \leqslant \log\operatorname{Tr}(e^H(A^p\#_\alpha B^p)^{1/p}). \qquad \square$$

Remarks. For $\alpha = p = 1$, Theorem 4.3 (ii) reduces to the case $B_1 = I_n$ of the Theorem 3.1 of Ando and Hiai. If $\alpha = 1$ and $L = 0$, then $(e^{pH}\#_\alpha e^{pL})^{1/p} = I_n$, $\operatorname{Tr}(\hat{S}(e^{pH}|e^{pL})\,e^{(1-p)H}) = -p\operatorname{Tr}(e^H H)$, and so Theorem 4.3 (iii) reduces to the Peierls–Bogoliubov inequality.

Final comments. Let H and L be Hermitian matrices, $0 \leqslant \alpha \leqslant 1$ and $p > 0$. Considering $A = e^H$ and $B = e^L$ in (7), we find that

$$\operatorname{Tr}(e^H)\exp\frac{\alpha\operatorname{Tr}(e^H(L-H))}{\operatorname{Tr}(e^H)} \tag{24}$$

is an upper bound to

$$\operatorname{Tr}(e^H)\exp\frac{\alpha\operatorname{Tr}(\hat{S}(e^{pH}|e^{pL})\,e^{(1-p)H})}{p\operatorname{Tr}(e^H)}. \tag{25}$$

Taking $K = -H$ in Theorem 4.3 (iii), another upper bound to (25) is

$$\operatorname{Tr}(e^{pH}\#_\alpha e^{pL})^{1/p}. \tag{26}$$

Moreover, as follows from the Peierls–Bogoliubov inequality, (24) is a lower bound to $\operatorname{Tr}(e^{(1-\alpha)H+\alpha L})$. By the complemented Golden–Thompson inequality, (26) is also a lower bound to $\operatorname{Tr}(e^{(1-\alpha)H+\alpha L})$. So, the comparision of these two bounds is of interest.

For each $0 < \alpha_0 \leqslant 1$, and H, L Hermitian such that $L - H$ is a non-scalar matrix, there exists $p_0 > 0$ (depending on H, L, α_0) such that the inequality holds

$$\operatorname{Tr}(e^H)\exp\frac{\alpha\operatorname{Tr}(e^H(L-H))}{\operatorname{Tr}(e^H)} < \operatorname{Tr}(e^{pH}\#_\alpha e^{pL})^{1/p}, \tag{27}$$

for $0 < p \leqslant p_0$ and $\alpha_0 \leqslant \alpha \leqslant 1$. If $L - H$ is a scalar matrix, equality occurs in (27), for all $p > 0$.

Indeed, suppose that for any $\epsilon > 0$ there exists p such that $0 < p \leqslant \epsilon$ and

$$\operatorname{Tr}(\mathrm{e}^H) \exp \frac{\alpha \operatorname{Tr}(\mathrm{e}^H(L-H))}{\operatorname{Tr}(\mathrm{e}^H)} \geqslant \operatorname{Tr}(\mathrm{e}^{pH} \#_\alpha \mathrm{e}^{pL})^{1/p}. \tag{28}$$

Let $A = \{p : (28) \text{ holds}\}$. This set has 0 as an accumulation point. Consider a sequence in A converging to 0. Taking limits as $p \downarrow 0$ in that sequence and recalling (9), we get

$$\operatorname{Tr}(\mathrm{e}^H) \exp \frac{\alpha \operatorname{Tr}(\mathrm{e}^H(L-H))}{\operatorname{Tr}(\mathrm{e}^H)} \geqslant \operatorname{Tr}(\mathrm{e}^{(1-\alpha)H+\alpha L}). \tag{29}$$

Since the Peierls–Bogoliubov inequality gives the previous inequality in reversed order, only the equality is possible in (29). The characterization of the equality condition in the Peierls–Bogoliubov inequality implies that $L - H$ is a scalar matrix, a contradiction. This proves that there exists $p_0 > 0$ such that (27) is verified. Now, let $L - H$ be a scalar matrix λI_n. For $p > 0$, both the left and the right hand side of (27) reduce to $\mathrm{e}^{\alpha\lambda} \operatorname{Tr}(\mathrm{e}^H)$, and so the equality occurs.

In conclusion, for each $0 < \alpha_0 \leqslant 1$, there exists $p_0 > 0$ such that the following inequalities hold

$$\begin{aligned}\operatorname{Tr}(\mathrm{e}^H) \exp \frac{\alpha \operatorname{Tr}(\hat{S}(\mathrm{e}^{pH}|\mathrm{e}^{pL})\, \mathrm{e}^{(1-p)H})}{p \operatorname{Tr}(\mathrm{e}^H)} &\leqslant \operatorname{Tr}(\mathrm{e}^H) \exp \frac{\alpha \operatorname{Tr}(\mathrm{e}^H(L-H))}{\operatorname{Tr}(\mathrm{e}^H)} \\ &\leqslant \operatorname{Tr}(\mathrm{e}^{pH} \#_\alpha \mathrm{e}^{pL})^{1/p} \\ &\leqslant \operatorname{Tr}(\mathrm{e}^{(1-\alpha)H+\alpha L}),\end{aligned}$$

for $0 < p \leqslant p_0$ and $\alpha_0 \leqslant \alpha \leqslant 1$, occurring equality when $L - H$ is a scalar matrix.

References

[1] T. Ando, F. Hiai, Log-majorization and complementary Golden–Thompson type inequalities, Linear Algebra Appl. 197 (1994) 113–131.

[2] T. Ando, F. Hiai, Hölder type inequalities for matrices, Math. Ineq. Appl. 1 (1998) 1–30.

[3] H. Araki, On an inequality of Lieb and Thirring, Lett. Math. Phys. 19 (1990) 167–170.

[4] N. Bebiano, J. Providência, Numerical ranges in physics, Linear Multilinear Algebra 43 (1998) 327–337.

[5] N. Bebiano, R. Lemos, J. Providência, Matricial inequalities in statistical mechanics, Linear Algebra Appl. 376 (2004) 265–273.

[6] V.P. Belavkin, P. Staszewski, C^*-algebraic generalization of relative entropy and entropy, Ann. Inst. H. Poincaré Sect. A 37 (1982) 51–58.

[7] J.I. Fujii, E. Kamei, Relative operator entropy in noncomutative information theory, Math. Japon. 34 (1989) 341–348.

[8] T. Furuta, $A \geqslant B \geqslant 0$ ensures $(B^r A^p B^r)^{1/q} \geqslant B^{(p+2r)/q}$, for $r \geqslant 0$, $p \geqslant 0$, $q \geqslant 1$ with $(1+2r)q \geqslant p+2r$, Proc. Amer. Math. Soc. 101 (1987) 85–88.

[9] S. Golden, Lower bounds for the Helmholtz function, Phys. Rev. 137 (1965) B1127–B1128.

[10] F. Hiai, D. Petz, The Golden–Thompson trace inequality is complemented, Linear Algebra Appl. 181 (1993) 153–185.

[11] F. Hiai, Equality cases in matrix norm inequalities of Golden–Thompson type, Linear Multilinear Algebra 35 (1994) 239–249.
[12] F. Kubo, T. Ando, Means of positive linear operators, Math. Ann. 246 (1980) 205–224.
[13] S. Kullback, R.A. Leibler, On information and sufficiency, Ann. Math. Statist. 22 (1951) 79–86.
[14] L.D. Landau, E.M. Lifshitz, Statistical Physics, Adison-Wesley Publishing Company Inc., Massachusetts, 1958.
[15] A.W. Marshall, I. Olkin, Inequalities: Theory of Majorization and its Applications, Academic, New York, 1979.
[16] M. Nakamura, H. Umegaki, A note on entropy for operator algebras, Proc. Jap. Acad. 37 (1961) 149–154.
[17] C.E. Shannon, The Mathematical Theory of Communication, University of Illinois Press, Urbana, 1949.
[18] C.J. Thompson, Inequality with applications in statistical mechanics, J. Math. Phys. 6 (1965) 1812–1813.
[19] H. Umegaki, Condition expectation in an operator algebra IV, Kodai Math. Sem. Rep. 14 (1962) 59–85.